

# Scalable Accelerated Intelligent Charging Strategy Recommendation for Electric Vehicles Based on Deep Q-Networks

Xianhao Shen<sup>1</sup>, Zhen Wu<sup>2</sup>, Yexin Zhang<sup>3</sup>, Shaohua Niu<sup>4\*</sup>

College of Information Science and Engineering, Guilin University of Technology, Guilin, China<sup>1, 2, 3</sup>

Guangxi Laboratory of Embedded Technology and Intelligent System, Guilin University of Technology, Guilin, China<sup>1, 2, 3</sup>

School of Mechanical and Electrical Engineering, Beijing Institute of Technology, Beijing, China<sup>4</sup>

**Abstract**—With the rapid development of electric vehicles, their charging strategies significantly impact the overall power grid. Solving the spatiotemporal scheduling problem of vehicle charging has become a hot research topic. This paper focuses on recommending suitable charging stations for electric vehicles and proposes a scalable accelerated intelligent charging strategy recommendation algorithm based on Deep Q-Networks (DQN). The strategy recommendation problem is formulated as a Markov decision process, where the continuous sequence of regional charging requests within a time slice is fed into the DQN network as the input state, enabling optimal charging strategy recommendations for each electric vehicle. The algorithm aims to maintain regional load balance while minimizing user waiting time. To enhance the algorithm's applicability, a scalable, accelerated charging strategy framework is further proposed, which incorporates information filtering and shared experience pool mechanisms to adapt to different expansion scenarios and expedite strategy iterations in new scenarios. Simulation results demonstrate that the proposed DQN-based strategy recommendation algorithm outperforms the shortest path-first strategy, and the scalable, accelerated charging strategy framework achieves a 64.3% improvement in iteration speed in new scenarios, which helps to reduce the cloud server load and saves overheads.

**Keywords**—Scalable acceleration; smart charging; Deep Q-network; Markov decision

## I. INTRODUCTION

In recent years, the global energy structure has slowly transitioned towards low-carbon resources, with low-carbon energy gradually gaining a higher share in the power sector. China has also announced its efforts to achieve carbon neutrality by 2060, which will stimulate the development of the renewable energy industry in the country. According to the Renewable Energy Market Report 2023 published by the International Energy Agency, the global installed capacity of renewable energy saw an increase of over 50% in 2023 compared to the previous year, marking the most significant annual increment since 1999. By the end of 2023, the number of new energy vehicles in China reached 20.41 million, with 6.278 million charging piles available, resulting in an electric vehicle (EV) to charging pile ratio of approximately 3.3:1. The rapid growth of electric vehicles has led to an explosive demand for charging infrastructure, presenting both new challenges and opportunities for the power grid. In addition, in

the face of a vast domestic user group, the existing charging stations in cities are gradually overburdened due to their sparse and uneven distribution, resulting in a severe mismatch between the current rate of new charging piles in China and the growth rate of new EV sales, and an urgent need for construction. This brings new challenges and opportunities for the intelligent charging strategy for electric vehicles.

Existing research has mainly focused on two aspects: energy storage scheduling at charging stations [1-4] and recommendation of charging strategies for electric vehicles [5-8]. Energy storage scheduling involves storing electrical energy generated by photovoltaic power generation [9-11] and managing cross-temporal energy dispatch to allocate electricity across different charging scenarios, mitigating sustained load pressure on the power grid [12]. In reference [13], a cross-temporal scheduling model integrating photovoltaic power and energy storage systems was constructed. It stored electricity during periods of low power consumption and released it during peak periods to meet changing demands. However, such cross-temporal energy scheduling algorithms rely on accurate energy usage prediction and suffer from limited energy storage efficiency and high costs. Regarding the research on the recommendation of charging strategies for electric vehicles, reference [14] developed a data-driven framework for energy prediction and utilized dynamic programming algorithms to seek optimal charging strategies. However, data-driven approaches become increasingly ineffective as the volume of data grows. Reference [15] proposes a strategy for the localization and route planning of public charging infrastructure for logistics companies based on a two-tier scheme. A two-tier genetic algorithm is used to derive the optimal routing and charging plan, and a simulated annealing descent algorithm is used to select charging station locations. The proposed method is tested and compared with a meta-heuristic approach using a benchmark instance with charging stations. Reference [16] proposes a nonlinear integer programming model with multiple objectives, including minimizing the average daily acquisition and charging costs of the electric bus routes, minimizing the time cost of waiting for charging of the electric buses and maximizing the charging revenues of the electric buses to synergistically realize the vehicle types allowed to be charged in each time window, the daily service journeys and charging journeys allocated to each electric bus. Subsequently, an algorithm was developed to

solve the formulated optimization model by combining enumeration with branching and pricing to solve the nonlinear problem. Reference [17] explored ordered charging strategies for electric vehicles using Monte Carlo algorithms, but the probabilistic nature of Monte Carlo algorithms introduces uncertainties in accurately assessing the quality of strategies. Reference [18] proposed a decision framework for charging and repositioning agent-based Shared Autonomous Electric Vehicles (SAEVs) fleets, which adjusts charging before expected demand, spatially and temporally dispersing the demand to reduce peak loads on the power grid and minimizes anticipated costs for operators. However, this framework does not consider the temporal evolution of SAEV demand and Electric Vehicle Charging Station (EVCS) supply or the cost of electricity, as its objective function only seeks to minimize response time rather than balancing charging frequency and response time.

The recommendation of charging strategies for electric vehicles is essentially a temporal scheduling problem [19], but numerous uncertain factors complicate the problem. With the rapid development of reinforcement learning, Markov decision models are well suited for charging strategy recommendations. In reference [20], a novel Markov decision process was constructed, dividing all connected electric vehicles into groups at each time step based on their charging priorities. Reinforcement learning agents were then employed to determine the charging proportions for each group of vehicles during each time interval. However, the arrival time and battery level of each electric vehicle at the charging station must be known to allocate it to a priority group. In reference [21], a graph reinforcement learning-based representation method integrates multi-dimensional information from charging stations, traffic nodes, and grid buses into a graph using feature projections. Graph convolution of coupled system states can then be implemented to facilitate environment perception. In reference [22], a novel multi-agent mean-field hierarchical reinforcement learning (MFHRL) framework was proposed to provide proactive charging and relocation advice for electric taxi drivers, maximizing the long-term cumulative rewards of their orders. The framework employed hierarchical reinforcement learning, with the manager setting goals that inherently guide the decision-making of workers, who receive rewards for following these goals. The integration of each level in the two hierarchies with mean-field approximation was carried out to incorporate the mutual influence of agents in decision-making, enabling finer temporal resolution at short intervals. In reference [23], an incentive demand response model was proposed, analyzing user behavior through reinforcement learning and subsequently guiding users to select periods with sufficient power supply. However, this approach only addresses the temporal scheduling problem, while the spatial scheduling problem remains unresolved. In reference [24], a multi-agent spatiotemporal reinforcement learning approach was introduced, altering the charging decision of electric vehicles by simulating future competitive environments using a delayed access policy. Reference [25] employed neural networks as function approximators to model user demands, training a central agent to develop charging plans for electric vehicles. None of these spatio-temporal

scheduling strategies discusses the variability of the actual environment.

Considering that China is in a period of development of charging infrastructure construction, the number of charging stations in the region is increasing, and the expandability of the scheduling strategy in the actual operation process is particularly important. Existing recommendation studies seldom consider the stability of the electric power system and the load balance while adapting to the changing environment, resulting in the charging recommendation strategy having a high maintenance and upgrading cost, and the strategy's practicality is poor.

The main contributions of this paper are mainly as follows:

1) To address the spatiotemporal scheduling issues in traditional electric vehicle charging strategies, a smart charging strategy recommendation algorithm based on Deep Q Network (DQN) is proposed. In this approach, the charging requests within a time slot are treated as a continuous sequence of charging request states and fed into the DQN network to generate optimal charging strategy recommendations for each electric vehicle.

2) To enhance the applicability of the proposed algorithm, an expandable and accelerated regional charging strategy recommendation algorithm framework is introduced. This framework utilizes a shared experience pool strategy to store strategy experiences from different regions. When a new region is added, the framework prioritizes training using experiences stored in the shared experience pool. At the same time, new experiences are stored in the experience pool of the new region. This significantly reduces the training iteration time of the model. Additionally, leveraging the experiences in the shared experience pool allows the model to converge faster and better fit the charging patterns of new regions.

The overall structure of this paper is as follows. Section II provides an introduction to the smart charging strategy recommendation model for electric vehicles based on Deep Q Network (DQN). In Section III, an expandable and accelerated regional charging strategy recommendation algorithm network framework is proposed. Section IV discusses the simulation results of the algorithm. Finally, Section V presents the conclusions and future directions for further work.

## II. SMART CHARGING STRATEGY RECOMMENDATION ALGORITHM FOR ELECTRIC VEHICLES BASED ON DQN

### A. Basic DQN Concepts

Deep Reinforcement Learning (DRL) [26] is a combination of Deep Learning (DL)[27] and Reinforcement Learning (RL)[28], which retains the ability of RL to solve policy problems. It involves the continuous interaction between an individual agent and an unknown environment, where the agent takes relevant control actions to maximize its future rewards. In theory, the value function can compute the reward value for any state and action, using methods such as Q-learning [29]. The Q-learning approach stores the state-action pairs and their corresponding rewards in a table, and when the state transitions to an environment corresponding to a table entry, the action's

reward value is obtained through table lookup. However, when there are a large number of states and actions, the computation or query time for the value or lookup function significantly increases.

The key difference between DQN and RL lies in the use of neural networks to approximate the agent's value function. Specifically, the state,  $s$ , is used as input to the neural network, and the output is the value  $Q(s, a)$  and its corresponding action,  $a$ . The Greedy( $s, a$ ) function is then combined with  $Q_{\max}(Q, a)$  to select the best value action while maintaining a certain level of exploration. DQN calculates the current action value in a manner similar to Q-learning, using the difference between it and the output of the value neural network as the loss value. This loss value is then passed into the loss function for iterative learning. During the iterative learning process, the insertion of memories from the experience pool facilitates mixed learning, resulting in a more efficient update of the neural network.

$$y_t = \begin{cases} R_t & \text{##end} \\ R_t + \gamma \max_{a_{t-1}} Q_{t-1}(s_{t-1}, a_{t-1}) & \text{#####not end#####} \end{cases} \quad (1)$$

$$L_\theta = E(y_t - Q_t(s_t, a_t, \theta))^2 \quad (2)$$

Eq. (1) and Eq. (2) represent the target value,  $y_t$  and the reward value,  $R_t$ , respectively, at time step  $t$ . The learning discount rate is denoted as  $\gamma$ , and  $Q_{t-1}(s_{t-1}, a_{t-1})$  represents the value network value at time step  $t - 1$ .  $L_\theta$  corresponds to the value of the loss function.

### B. A DQN-based Recommendation Model for Smart Charging Strategies for Electric Vehicles

This study primarily focuses on providing charging strategy recommendations for electric vehicles (EVs) at public charging stations. As illustrated in Fig. 1, EVs with charging demands within a designated area send their charging requests to a central processor. They also transmit their specific vehicle information, including current battery level and location, to the central processor. The central processor collects all the charging requests from EVs within the same time period in the area, forming a temporal sequence of charging requests. This sequence serves as the input to the Deep Q-Network (DQN) for generating optimal charging strategy recommendations for all EVs within a time slice. Considering the timeliness of charging strategy recommendations, the study employs time slicing by dividing each minute into 60 time slices, with each time slice representing 1 second. Within a time slice, the processor composes timing input vectors from the states of all requests combined with the load conditions of the charging station and charging pile information, etc., and makes the correct strategy decision for the EV through a deep reinforcement learning model to guide the EV to complete the charging, which satisfies the need to maintain the load of the regional power grid while shortening the user's waiting time.

The recommendation of charging strategies for electric vehicles (EVs) can be viewed as a Markov decision process, which involves coordinating the interaction between EVs and the regional charging environment. The goal is to guide each EV to make informed decisions regarding charging strategies

while minimizing user waiting time and the load on the regional power grid. However, treating each individual EV as the main agent does not satisfy the continuity of the state space in the Markov decision process. Therefore, a time slicing approach is adopted, where all charging requests from EVs within a time slice in the region are sorted based on their submission time, forming a continuous state space for the regional charging requests. As shown in Fig. 2 and described by Eq. (3) and Eq. (4), when the agent  $C_1$  submits request  $q_1$ , its current location state  $Qarea$ , state of charge  $Qsoc$ , and the location information of each charging station  $Larea$  are combined to form the overall state  $s_{q_1}$ . Subsequently, the state transitions from  $s_{q_1}$  to  $s_{q_2}$  as the vehicle  $C_2$  request is processed. The collection of all request states  $s_{q_m}$  within the time slice  $t_1$  forms the aggregate state  $s_{t_n}$ , which serves as the input to the DQN network for training in a single episode.

$$s_{q_m} = (Qarea_{q_m}, Qsoc_{q_m}, Larea_{q_m}) \quad (3)$$

$$s_{t_n} = \{s_{q_1}, s_{q_2}, \dots, s_{q_m}\} \quad (4)$$

The DQN action space employed in this study corresponds to the selection of charging stations, where EVs continuously make decisions on charging stations within a time slice, and the action space is the same for all requests. As shown in Fig. 3, the action space corresponds to different discrete charging stations. The agent can choose from the following four actions: 1、Charging station 1, 2、Charging station 2, 3、Charging station 3, and 4、Charging station 4. These charging stations are randomly distributed, and their initial charging states are also randomized. The agent is trained in various stochastic environments to cope with challenges in real-world settings.



Fig. 1. Scenario of the use of DQN-based recommendation model for smart charging strategy for electric vehicles.

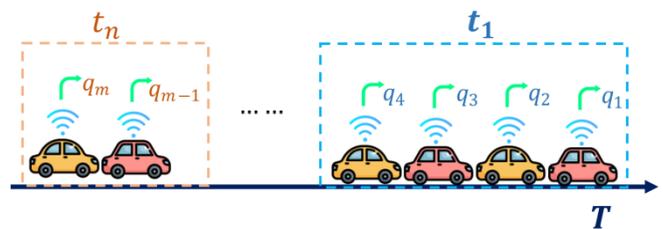


Fig. 2. Time slice model diagram.

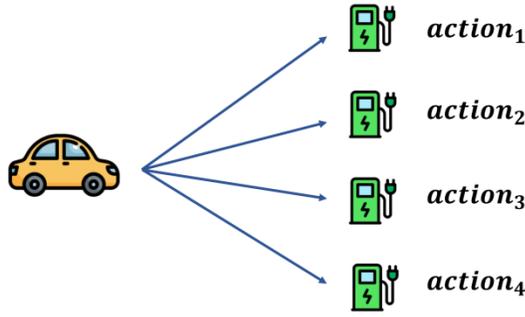


Fig. 3. Action space diagram.

Rewards provide direct or delayed feedback to the agent's decisions, enabling the agent to continually update its decisions to maximize the rewards. Rewards quantify higher-level objectives in multi-agent reinforcement learning. Specifically, in the context of electric vehicle charging strategies, the reward is set as a composite reward to expedite the training iteration of the intelligent agent. Upon making a decision regarding the request  $q_t$ , the intelligent agent receives the reward functions  $r_{dis}(s_{q_m}, a_{q_m})$ ,  $r_{wait}(s_{q_m}, a_{q_m})$ , and  $r_{load}(s_{q_m}, a_{q_m})$  as defined in Eq. (5) and Eq. (6), respectively:

$$\left\{ \begin{array}{l} r_{dis}(s_{q_m}, a_{q_m}) = \begin{cases} -\frac{dis_{range}}{dis_{min}} dis_{range} + dis_{min} < mile_{min} \\ -dis_{range} dis_{range} + dis_{min} > mile_{min} \end{cases} \\ r_{wait}(s_{q_m}, a_{q_m}) = -t_{wait} \\ r_{load}(s_{q_m}, a_{q_m}) = \begin{cases} \sum^i n_{charge} \frac{1}{n_{rest}} \\ \sum^i n_{charge} \frac{1}{n_{wait}} \end{cases} \end{array} \right. \quad (5)$$

$$R_t(st_i, a_t) = \partial r_{dis}(st_i, a_t) + \beta r_{wait}(st_i, a_t) + \delta r_{load}(st_i, a_t) \quad (6)$$

The variables in the equation are defined as follows:  $dis_{min}$  represents the shortest distance to the charging station,  $dis_{error}$  denotes the difference between the selected station and the shortest station in terms of distance.  $mile_{min}$  represents the minimum remaining mileage of the vehicle based on its current condition.  $t_{wait}$  represents the charging waiting time.  $n_{charge}$  corresponds to the sequence number of the charging station.  $n_{rest}$  and  $n_{wait}$  respectively indicate the number of available charging piles at the currently selected station and the number of vehicles queuing for charging.  $s_{nc}$  represents the current charging status of the station, which collectively determines the overall load of the region. Finally,  $\partial$ ,  $\beta$ , and  $\delta$  are discount factors.

Observation value: To give the central processor a better grasp of the global information, an observation value is set for each intelligent body. They are set as shown in Eq. (7).

$$obs_m = \{Qarea_{q_m}, fpile_{q_m}, load_{q_m}\} \quad (7)$$

In the equation,  $s_m$  represents the state set composed of the position status and battery status of all electric vehicles within the current region.  $f_m$  denotes the number of available charging piles in the region, while  $l_m$  represents the total load of the region.

### C. Reinforcing the Learning Process

The recommended smart charging strategy for electric vehicles based on DQN is illustrated in Fig. 4. The process

begins by initializing the experience replay buffer, neural network parameters, and the initial state denoted as  $s$  in the DRL model. Subsequently, the states of all electric vehicles within the region are collected to form a state set. The charging policy network and the charging value network are separately utilized to obtain the actual reward  $r$ , value network reward  $r'$ , next state  $s'$ , and action  $a$ . These parameters are then stored in the experience replay buffer. At irregular intervals, parameters are randomly sampled from the experience replay buffer and added to the EV state set for training. Following this, the value network reward  $r'$  and the actual reward  $r$  are input into the loss function to train the charging value network. The EV state  $s$  is updated to  $s'$  and the EV state set is updated iteratively until the current training round is completed. Finally, the next LSTM model predicts the EV state, and this process continues until the training is completed, resulting in the output of the trained charging value network model.

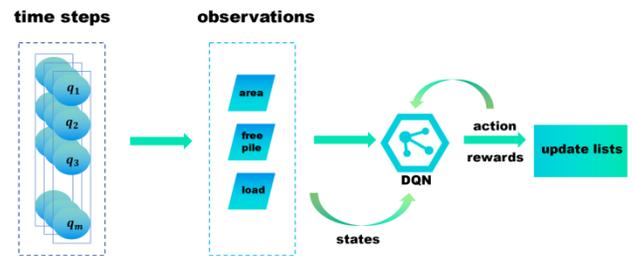


Fig. 4. Diagram of recommended smart charging strategies for electric vehicles.

## III. AN ALGORITHMIC FRAMEWORK FOR RECOMMENDING REGIONAL CHARGING STRATEGIES WITH SCALABLE ACCELERATION

### A. Framework Background

Currently, electric vehicles are undergoing an incredible and rapid development, leading to a continuous increase in charging demand. To alleviate the pressure on charging load, many regions have started constructing new charging stations. However, existing charging strategies [30-32] have not addressed their scalability. Adding a new charging area and starting the training of charging recommendation strategies from scratch undoubtedly incurs additional costs. Therefore, this paper proposes a scalable and accelerated framework for regional charging strategy recommendation algorithm.

### B. Framework Scenario Analysis

The individual charging station information within a single region is presented in Fig. 5 and Fig. 6, including the operational status, available quantity of charging piles, and specific locations of the charging stations. Initially, the information filtering layer is employed to select the information from the  $n$  closest charging stations to the charging-requesting vehicle, forming a new tuple of charging station information features with a length of  $n$ . The specific value of  $n$  will be described in detail in the experimental section. Subsequently, the new tuple of features is input into the DQN network for training, ultimately providing policy recommendations. The initial input states, decisions, rewards, and other parameters for each policy recommendation are stored in the network's own experience replay buffer and the

shared experience pool of the extended framework. During the training of policy recommendation across multiple regions, when updating the policy value network, random sampling from the shared experience pool is incorporated to achieve experience sharing. This facilitates accelerated training when new regions join the shared experience pool, effectively avoiding the issue of random recommendations due to insufficient initial experience pool capacity. Furthermore, the self-experience replay buffer is continually improved during the training process. Once its capacity is full and construction is completed, the framework utilizes its own experience replay buffer. Next, the applicability of this framework will be discussed based on three extended scenarios [33].

Scenario 1: Addition of new charging piles within the region. The purpose of this algorithm is to recommend the optimal charging station. Within the algorithm environment, there is a queue of information regarding available charging piles at the charging stations. When new charging piles are added to a charging station, it simply increases the count of available charging piles, without affecting the functionality of the algorithm.

Scenario 2: Addition of new charging stations within the region. The first layer of the proposed recommendation algorithm framework filters the information of all charging stations within the region. It retains a tuple of information features with a length of  $n$  ensuring that the input dimension of the DQN network remains consistent. This, in turn, guarantees consistency in the action space dimension of the DQN network. Specifically, when a vehicle makes a request, the DQN network takes a filtered queue of  $n$  nearest charging station information as input and ultimately provides policy recommendations among these  $n$  charging stations for the vehicle.

Scenario 3: Addition of a new charging area. The shared experience pool within the proposed recommendation algorithm framework is designed to address this scenario. The new region can directly utilize the shared experience pool to accelerate training, continuously accumulate and improve its own experience replay buffer, and eventually develop its specific charging strategy.

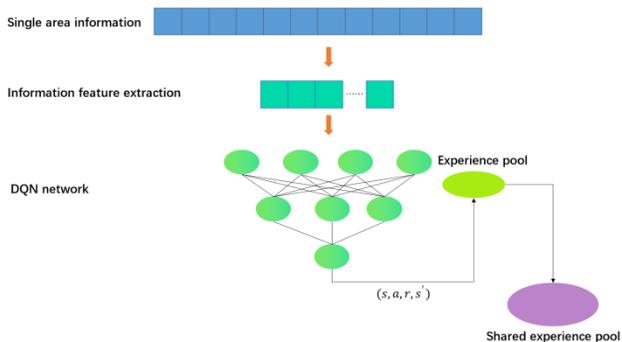


Fig. 5. Framework of the scalable regional charging policy recommendation algorithm.

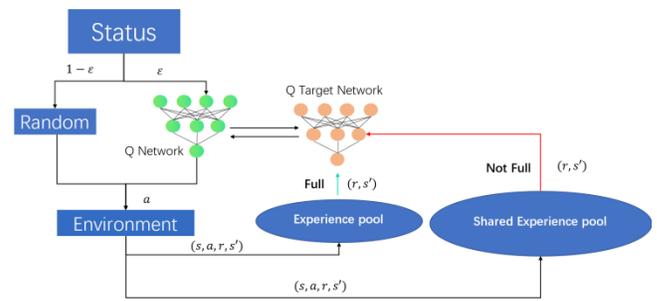


Fig. 6. Flow chart of the shared experience pool.

#### IV. ALGORITHM ANALYSIS

##### A. DQN-based Algorithm for Recommending Smart Charging Strategies for Electric Vehicles

1) *Experiment description:* This experiment primarily simulates the decision-making behavior of electric vehicles in a region regarding public charging stations. The region is set to a size of 2000\*2000 grids, and at the beginning of each experimental round, the coordinates of the charging stations within the region, as well as the positions and coordinates of the charging requests within the region, are randomly initialized. The coordinates are used to simulate real-world latitude and longitude.

In this experiment, a comparison will be made between the DQN-based intelligent charging strategy recommendation algorithm and the nearest distance-first strategy in terms of specific performance metrics such as average charging waiting time and average regional load. The experiment involves storing the location information of the region's charging stations on a server and simulating the application scenarios of the scalable regional charging strategy recommendation algorithm framework through local-server interactions.

2) *Parameter setting:* To validate the proposed algorithm, the following experiments were conducted in the simulation environment as shown in Table I. In this algorithm, the batch size of 32 was selected for each training iteration. The learning rate  $lr$  of the DQN network was set to 0.01, the exploration-exploitation trade-off rate  $\epsilon$  was set to 0.9, and the discount factor  $\gamma$  for the policy was set to 0.9. The experience replay buffer size was set to 100,000, and the target network was updated every 100 iterations .

TABLE I. SIMULATION ENVIRONMENT CONFIGURATION TABLE

Configuration of experiment	Specific Parameters
CPU	Intel Core i7-11700K@ 5.0GHz
GPU	NVIDIA GeForce GTX 1660 Super
Memory	16GB
Operating system	Windows10
Programming environment	python3.7、pytorch 1.10.2+cu102
Reinforcement environment	learning gym0.10.5

3) *Analysis of results:* After 5000 iterations of training, as shown in Fig. 7, where the first 1000 iterations were used for the experience replay buffer population, the reward values for the DQN-based intelligent charging strategy recommendation algorithm converged to approximately -3.5. To demonstrate the performance of the proposed algorithm, we will now discuss the simulation results in detail. Fig. 8(a) presents the load situation of the charging strategy recommendation algorithm based on DQN, while Fig. 8(b) shows the load situation for the nearest distance-first strategy. It can be observed that with the increase in time steps, our proposed algorithm exhibits some fluctuations. However, it shows an overall decreasing trend, significantly different from the nearest distance-first strategy. Based on calculations, the average load per step for the DQN-based charging strategy recommendation algorithm is 1.14, while for the nearest distance-first strategy, it is 1.20, resulting in an improvement of approximately 5.0%.

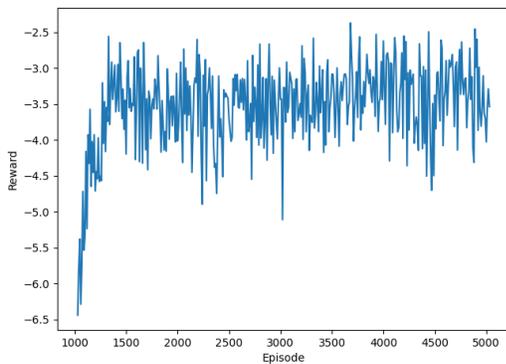


Fig. 7. Iteration diagram of the DQN-based charging policy recommendation algorithm model.

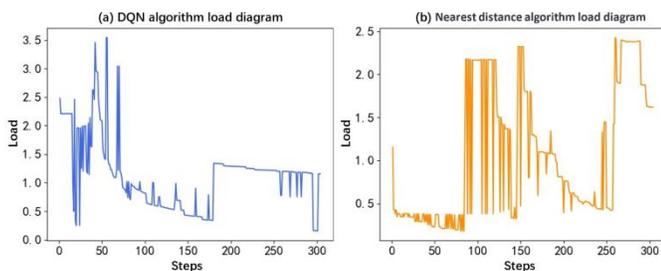


Fig. 8. DQN based intelligent charging strategy recommendation algorithm for electric vehicles and the nearest distance recommended load comparison diagram.

In terms of waiting time, as shown in Fig. 9, where Fig. 9(a) represents the DQN-based electric vehicle charging strategy recommendation algorithm's ability to make correct recommendations for immediate use when all charging stations in the area are initially vacant and to reasonably schedule charging plans even when all charging stations are under load in the latter part. In contrast, Fig. 9(b) depicts the nearest distance-first strategy, which fails to make optimal charging plan arrangements from the beginning. Based on calculations, the average waiting time per step for the DQN-based charging strategy recommendation algorithm is 1.75 ms, while for the

nearest distance-first strategy, it is 1.91 ms, resulting in an improvement of approximately 8.37%. It can be observed that the DQN-based charging strategy recommendation algorithm proposed in this paper not only maintains balanced area loads but also significantly reduces users' waiting time, which helps alleviate user anxiety and enhances the user experience.

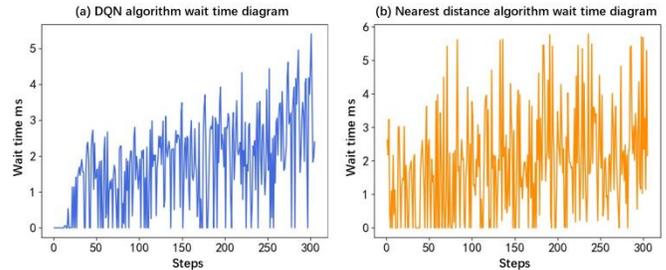


Fig. 9. DQN based intelligent charging strategy recommendation algorithm for electric vehicles and the nearest distance recommended waiting time comparison diagram.

In addition to waiting time, the distance to the recommended charging station is also a criterion for measuring the algorithm's accuracy. Fig. 10(a) and 10(b) below represent the DQN-based electric vehicle charging strategy recommendation algorithm and the nearest distance recommendation algorithm, respectively. From the figures, it can be observed that the DQN-based recommendation algorithm is primarily consistent with the nearest distance recommendation algorithm. Out of 300 testing steps, the DQN-based recommendation algorithm recommended the nearest distance priority in 121 cases. In contrast, in the remaining 179 cases, it selected other charging stations to minimize total time.

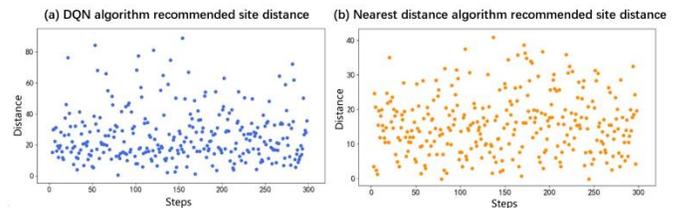


Fig. 10. DQN based intelligent charging strategy recommendation algorithm for electric vehicles and the nearest distance recommendation algorithm recommended site distance comparison diagram.

In terms of enablement, the experiments tested the DQN-based scalable EV smart charging policy recommendation algorithm model size of 1.21M with an average delay of 923ms, which has a strong real-time performance and can be applied to practical scenarios.

### B. Scalable Acceleration Algorithm for Electric Vehicle Charging Strategy Recommendation

1) *Experiment description:* This experiment mainly simulates the expansion charging scenario in the region, the iteration speed of the model will be verified separately, and the expandability in different scenarios.

2) *Parameter setting:* The experimental parameters of the network part of this experiment are consistent with the DQN strategy algorithm above, i.e. the number of samples selected for one training session is 32, the learning rate  $\text{lr}$  is 0.01, the

greed rate  $\epsilon$  is 0.9, the discount rate  $\gamma$  is 0.9, the experience pool size is set to 100000, and the target network following frequency is 100. The total experience pool size in the algorithm is set to 10,000,000 and the following frequency is 10,000.

3) *Analysis of results:* To cope with increasingly complex charging scenarios, the proposed scalable and accelerated electric vehicle charging strategy recommendation algorithm in this paper filters the charging station information within a single region through an information filtering layer. It forms a new charging station information feature tuple of length  $n$ , as illustrated in Fig. 11. In order to observe the impact of  $n$  on the waiting time in the recommendation algorithm, we conducted the following experiments, and the average waiting time was minimized when  $n$  was set to 9.

To validate the feasibility of the algorithm, this study conducted simulation experiments on the following scenarios based on practical application scenarios: Scenario 1: Adding new charging poles to charging stations within a region; Scenario 2: Adding new charging stations within a region; Scenario 3: Adding new charging regions; Scenario 4: Complex real-world scenarios. In Scenario 4, the number of experimental subjects in Scenario 3 was doubled. Specific parameters are shown in Table II.

Fig. 12 presents a comparison of the average waiting time for each scenario. Fig. 12(a) shows the original waiting time graph obtained from Experiment IV (A), while Fig. 12(b), Fig. 12(c), and Fig. 12(d) correspond to Scenario 1, Scenario 2, and Scenario 3, respectively. In Scenario 1, adding new charging poles within the region provides the algorithm with more choices, resulting in a significant decrease in the average waiting time to 0.59 ms. In Scenario 2, adding new charging stations slightly reduces the average waiting time to 0.96 ms. In Scenario 3, expanding the charging region results in a decreased average waiting time of 1.69 ms, representing improvements of 66.3%, 45.1%, and 3.4%, respectively. Thus, it can be concluded that the proposed scalable and accelerated electric vehicle charging scheduling recommendation algorithm remains applicable in complex scenarios, and its performance improves as the complexity of the scenarios increases.

In terms of the load aspect, as shown in Fig. 13, Fig. 13(a) represents the original load graph, Fig. 13(b) represents the load graph for Scenario 1 with an average load reduction of 0.8, Fig. 13(c) represents the load graph for Scenario 2 with an average load reduction of 0.84, and Fig. 13(d) represents the load graph for Scenario 3 with an average load reduction of 1.12. These reductions correspond to 29.8%, 26.3%, and 1.75% improvements, respectively. The average algorithm latency for each scenario is 889 ms, 893 ms, and 897 ms, representing reductions of 3.7%, 3.2%, and 2.8%, respectively. In conclusion, the proposed scalable and accelerated electric vehicle charging scheduling recommendation algorithm in this chapter reduces user waiting time while ensuring the stability of the regional load in complex scenarios. This contributes to better revenue generation for operators.

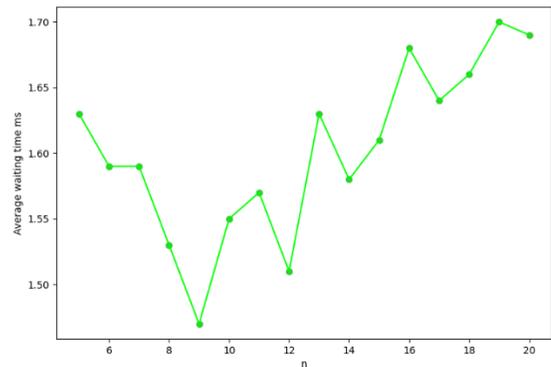


Fig. 11. Graph of results for feature tuple length  $n$ .

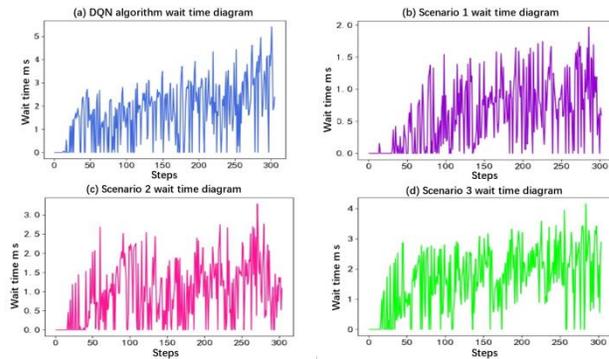


Fig. 12. Comparison of the average waiting time in each scenario.

TABLE II. COMPARISON TABLE OF SCENE PARAMETERS

Scene Name	The original scene	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Size of grid	2000*2000	2000*2000	2000*2000	2000*2000	6000*6000
Number of requests per unit time	300	300	300	300	300
Number of charging stations	20	20	40	20	40
Number of charging piles per charging station	20	40	20	20	40

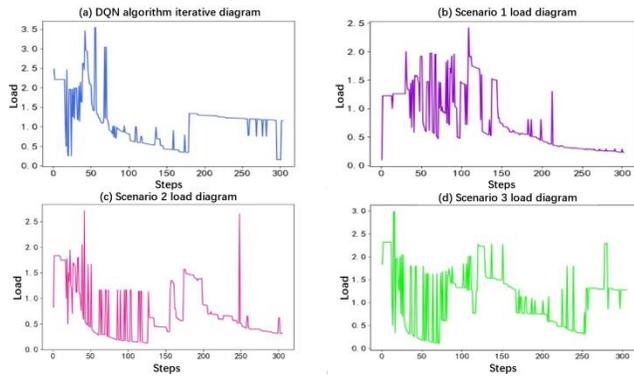


Fig. 13. Load comparison diagram of each scenario.

To validate the applicability of the proposed algorithm in real-world complex scenarios, we introduced increased complexity to the scenario parameters, and the results are shown in Fig. 14. Fig. 14(a) depicts the iteration graph of the charging strategy recommendation algorithm based on DQN, which converges after approximately 1400 iterations due to the need for experience pool storage. Fig. 14(b) represents the iteration graph for Scenario 3, while Fig. 14(c) corresponds to the iteration graph for Scenario 4. It is evident that compared to the DQN-based charging strategy recommendation algorithm, expanding the new region in training, as proposed in this study, using the shared experience pool approach eliminates the time required for storing the experience pool. Moreover, the experiences generated by the shared experience pool, compared to those randomly selected by DQN for action generation, are more practical and accelerate the fitting of model parameters, resulting in faster model iterations. In particular, the iteration speed is improved by 64.3% (500 iterations) and 67.8% (450 iterations) for Scenario 3 and Scenario 4, respectively. This significantly reduces the load on the cloud server and saves costs. Regarding waiting time, as shown in Fig. 15, Scenario 4, with the addition of more charging stations and charging piles and an expanded map area, offers users more choices, leading to a decrease in average waiting time compared to Scenario 3, reaching 0.51ms. Complex scenarios often accompany increased model execution time. However, in the simulated experiments of this algorithm in Scenario 4, the average algorithm latency remained relatively unchanged at 901ms, as mentioned earlier. The results demonstrate that as the complexity of the application scenarios increases, this algorithm can further accelerate the model iteration speed, reduce average waiting time for users, and maintain a consistent algorithm latency, showcasing its high applicability.

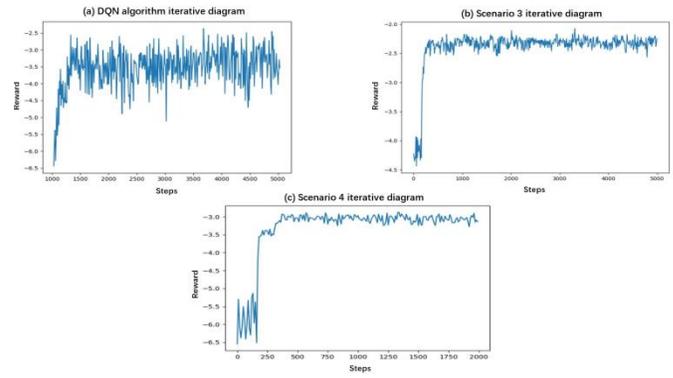


Fig. 14. Comparison diagram of training iteration speed of each scene.

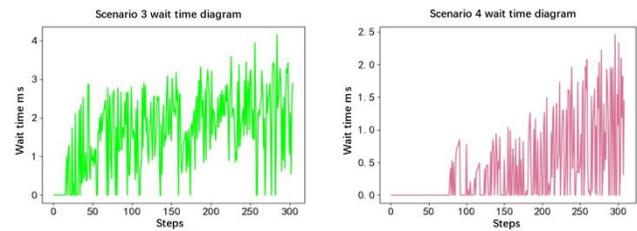


Fig. 15. Scenario 3 and Scenario 4 Waiting time comparison.

## V. CONCLUSION

This paper presents an intelligent electric vehicle (EV) charging strategy recommendation algorithm based on Deep Q-Network (DQN). The algorithm utilizes Markov modeling of user-requested charging events to formulate reasonable charging plans and effectively addresses the spatial scheduling issues in traditional EV charging strategies. Considering the rapid development of charging infrastructure construction in China, we propose a scalable and accelerated regional charging strategy recommendation algorithm framework. This framework not only adapts to increasingly complex and evolving charging scenarios but also maintains a consistent algorithm latency, further accelerating the iteration of the algorithm model. Experimental results show that the algorithm can improve the efficiency of charging strategy recommendation, charging waiting time, and charging demand response speed. In contrast, the expandable and accelerated charging strategy framework improves the iterative speed by 64.3% in new scenarios, which reduces the cloud server load and saves overheads. In future work, we will further refine the hardware implementation of the algorithm to realize a more efficient, precise, and practical charging strategy recommendation algorithm. This will provide superior, efficient, and convenient charging services for EVs, positively contributing to the development of innovative urban transportation.

## ACKNOWLEDGMENT

This research was partly funded by the National Natural Science Foundation of China (No. 62362017, No. 62341117).

## REFERENCES

- [1] M.S. Lipu Hossain, et al, "A review of controllers and optimizations based scheduling operation for battery energy storage system towards decarbonization in microgrid: Challenges and future directions," *Journal of Cleaner Production*, vol. 360, pp. 132188, Aug. 2022.
- [2] J. Massana, L. Burgas, S. Herraiz, J. Colomer, and C. Pous, "Multi-vector energy management system including scheduling electrolyser, electric vehicle charging station and other assets in a real scenario," *Journal of Cleaner Production*, vol. 380, pp. 134996, Dec. 2022.
- [3] G. Jianwei, G. Fangjie, Y. Yu, et al, "Configuration optimization and benefit allocation model of multi-park integrated energy systems considering electric vehicle charging station to assist services of shared energy storage power station," *Journal of Cleaner Production*, vol. 336, pp. 130381, Feb. 2022.
- [4] L. Desheng, Z. Adama, L. Jiantang, and Y. Hongtzer, "An energy management strategy with renewable energy and energy storage system for a large electric vehicle charging station," *eTransportation*, vol. 6, pp. 100076, 2020.
- [5] M.A. Hannan, M.S. Mollik, A.Q. Al-Shetwi, et al, "Vehicle to grid connected technologies and charging strategies: Operation, control, issues and recommendations," *Journal of Cleaner Production*, vol. 339, pp. 130587, Mar. 2022.
- [6] Z. Jiaqing, L. Jing, and Z. Xiaohui, "Charging navigation strategy for electric vehicles considering empty-loading ratio and dynamic electricity price," *Sustainable Energy, Grids and Networks*, vol. 34, pp. 100987, Jun. 2023.
- [7] P. Priyadarshan, K. Kazemzadeh, and B. Prateek, "Integration of charging behavior into infrastructure planning and management of electric vehicles: A systematic review and framework," *Sustainable Cities and Society*, vol. 88, pp. 104265, Jan. 2023.
- [8] M. S. Muhammad, Z. Shengxian, M. M. Hafiz, et al, "A study of charging-dispatch strategies and vehicle-to-grid technologies for electric vehicles in distribution networks," *Energy Reports*, vol. 9, PP. 1777-1806, Dec. 2023.
- [9] Y. Y. Kah, C. H. Hon, and K. J. Jirfi, "Solar Energy-Powered Battery Electric Vehicle charging stations: Current development and future prospect review," *Renewable and Sustainable Energy Reviews*, vol. 169, pp. 112862, Nov. 2022.
- [10] X. Fangzhou, C. Hongkun, L. Hao, and C. Lei, "Optimal planning of photovoltaic-storage fast charging station considering electric vehicle charging demand response," *Energy Reports*, vol. 8, pp. 399-412, Nov. 2022.
- [11] Z. Sixiang, W. Yachao, J. Zhenyu, H. Tianshu, and C. Fengming, "Research on emergency distribution optimization of mobile power for electric vehicle in photovoltaic-energy storage-charging supply chain under the energy blockchain," *Energy Reports*, vol. 8, pp. 6815-6825, Nov. 2022.
- [12] Y. Meng, Z. Lihui, Z. Zhenli, and W. Liwan, "Comprehensive benefits analysis of electric vehicle charging station integrated photovoltaic and energy storage," *Journal of Cleaner Production*, vol. 302, pp.126967, Jun. 2021.
- [13] K. Kouka, A. Masmoudi, A. Abdelkafi, and L. Krichen, "Dynamic energy management of an electric vehicle charging station using photovoltaic power," *Sustainable Energy, Grids and Networks*, vol. 24, p. 100402, Dec. 2020.
- [14] Z. Yi and M. Shir, "Data-driven optimal charging decision making for connected and automated electric vehicles: A personal usage scenario," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 37-58, Jan. 2018.
- [15] L. Jiale, L. Zhenbo, and W. Xuefei, "Public charging station localization and route planning of electric vehicles considering the operational strategy: A bi-level optimizing approach," *Sustainable Cities and Society*, vol. 87, pp. 104153, Dec. 2022.
- [16] J. Jinhua, B. Yiming, and W. Linhong, "Optimal electric bus fleet scheduling for a route with charging facility sharing," *Transportation Research Part C: Emerging Technologies*, vol. 147, pp. 104010, Feb. 2023.
- [17] L. Hongyan et al., "Research on orderly charging control strategy of electric vehicles based on Monte Carlo algorithm," *Electrical Appliance and Energy Efficiency Management Technology*, no. 05, pp. 59-64, 2021.
- [18] M. D. Dean, K. M. Gurumurthy, F. de Souza, J. Auld, and K. M. Kockelman, "Synergies between repositioning and charging strategies for shared autonomous electric vehicle fleets," *Transportation Research Part D: Transport and Environment*, vol. 108, pp. 103314, Jul. 2022.
- [19] W. Ning, T. Hangqi, Z. Shunbo, and L. Yuan, "Analysis of public acceptance of electric vehicle charging scheduling based on the technology acceptance model," *Energy*, vol. 258, pp. 124804, Nov. 2022.
- [20] N. Sadeghianpourhamami, J. Deleu, and C. Develder, "Definition and Evaluation of Model-Free Coordination of Electrical Vehicle Charging With Reinforcement Learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 203-214, Jan. 2020.
- [21] X. Peidong, Z. Jun, G. Tianlu, et al, "Real-time fast charging station recommendation for electric vehicles in coupled power-transportation networks: A graph reinforcement learning method," *International Journal of Electrical Power & Energy Systems*, vol. 141, pp. 108030, Oct. 2022.
- [22] E. Wang et al, "Joint Charging and Relocation Recommendation for E-Taxi Drivers via Multi-Agent Mean Field Hierarchical Reinforcement Learning," *IEEE Trans. Mob. Comput.*, vol. 21, no. 4, pp. 1274-1290, Apr. 2022.
- [23] L. Xin et al, "Orderly charging control of electric vehicles based on MDP and incentive demand response," *Journal of Electric Power Science and Technology*, vol. 36, no. 05, pp. 79-86, 2021.
- [24] W. Zhang et al, "Intelligent Electric Vehicle Charging Recommendation Based on Multi-Agent Reinforcement Learning," *WWW*, 2021.
- [25] F. Tuhnitz, N. Ebell, J. Schlund, and M. Pruckner, "Development and Evaluation of a Smart Charging Strategy for an Electric Vehicle Fleet Based on Reinforcement Learning," *Applied Energy*, vol. 285, p. 116382, Mar. 2021.
- [26] K. Arulkumaran, M.P. Deisenroth, M. Brundage, and A.A. Bharath, "Deep Reinforcement Learning: A Brief Survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26-38, Sep. 2017.
- [27] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.
- [28] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238-1274, Sep. 2013.
- [29] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach Learn*, vol. 8, no. 3, pp. 279-292, May 1992.
- [30] P. Xu, J. Zhang, T. Gao et al, "Real-time fast charging station recommendation for electric vehicles in coupled power-transportation networks: A graph reinforcement learning method," *International Journal of Electrical Power & Energy Systems*, vol.141, p. 2108030, Oct. 2022.
- [31] X. Qiang, C. Zhong, Z. Ziqi, H. Xueliang, and L. Xiaohui, "Route Planning and Charging Navigation Strategy for Electric Vehicles Based on Real-time Traffic Information and Grid Information," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 752, no. 1, p. 012011, 2020.
- [32] J. Zhong, N. Yang, X. Zhang, and J. Liu, "A fast-charging navigation strategy for electric vehicles considering user time utility differences," *Sustainable Energy, Grids and Networks*, vol. 30, p. 100646, Jun. 2022.
- [33] S. Borozan, S. Giannelos, and G. Strbac, "Strategic network expansion planning with electric vehicle smart charging concepts as investment options," *Advances in Applied Energy*, vol. 5, p. 100077, Feb. 2022.